

533,669

## (12)特許協力条約に基づいて公開された国際出願

(19)世界知的所有権機関  
国際事務局(43)国際公開日  
2004年5月27日 (27.05.2004)

PCT

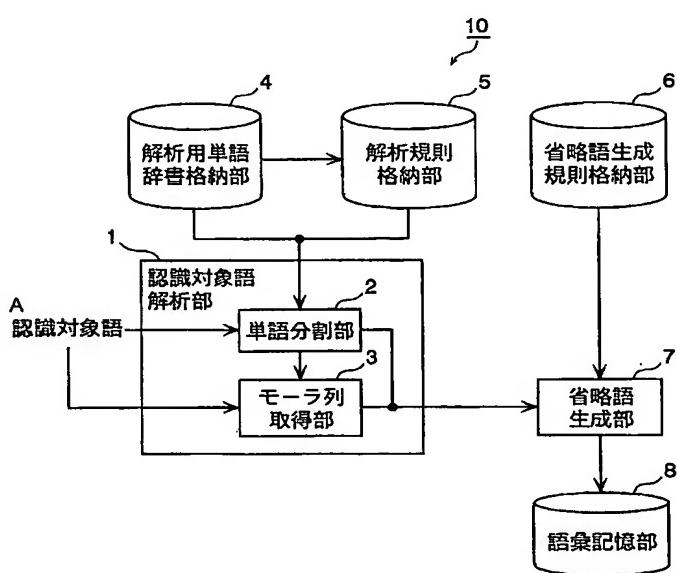
(10)国際公開番号  
WO 2004/044887 A1

- (51) 国際特許分類<sup>7</sup>: G10L 15/06, 15/18 (71) 出願人(米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府門真市大字門真1006番地 Osaka (JP).
- (21) 国際出願番号: PCT/JP2003/014168 (72) 発明者; および
- (22) 国際出願日: 2003年11月7日 (07.11.2003) (75) 発明者/出願人(米国についてのみ): 沖本 純幸 (OKI-MOTO, Yoshiyuki) [JP/JP]; 〒619-0224 京都府相楽郡木津町兜台2丁目2番1号 D-105号室 Kyoto (JP).
- (25) 国際出願の言語: 日本語 (74) 代理人: 新居 広守 (NII, Hiromori); 〒532-0011 大阪府大阪市淀川区西中島3丁目11番26号 新大阪末広センタービル3F 新居国際特許事務所内 Osaka (JP).
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願 2002-326503 2002年11月11日 (11.11.2002) JP

[続葉有]

(54) Title: SPEECH RECOGNITION DICTIONARY CREATION DEVICE AND SPEECH RECOGNITION DEVICE

(54) 発明の名称: 音声認識用辞書作成装置および音声認識装置



- 4...WORDS-TO-BE-ANALYZED DICTIONARY STORAGE SECTION
- 5...ANALYSIS RULE STORAGE SECTION
- 6...ABBREVIATED WORD CREATION RULE STORAGE SECTION
- 1...RECOGNITION OBJECT WORD ANALYSIS SECTION
- A...RECOGNITION OBJECT WORD
- 2...WORD SEPARATION SECTION
- 3...MORA STRING ACQUISITION SECTION
- 7...ABBREVIATED WORD CREATION SECTION
- 8...VOCABULARY STORAGE SECTION

納部(6)、構成単語ごとのモーラ列からモーラを取り出して連接することにより、1個以上のモーラからなる省略語の候補を生成するとともに、それら候補に対して、省略語生

(57) Abstract: A speech recognition dictionary creation device (10) can effectively create a speech recognition dictionary capable of recognizing even an abbreviated expression of a word with a high recognition ratio. The device includes: a word separation section (2) for dividing a recognition object speech consisting of one or more words into constituting words; a mora string acquisition section (3) for creating a mora string for each of the constituting words according to the reading of the constituting words separated; an abbreviated word creation rule storage section (6) for storing an abbreviated word creation rule using mora; an abbreviated word creation section (7) for taking a mora out of a mora string of each constituting word, concatenating them so as to create a candidate of an abbreviated word consisting of one or more moras and applying the abbreviated word creation rule to the candidates so as to create an abbreviated word; and vocabulary storage section (8) for storing the created abbreviated word together with the recognition object word as a speech recognition dictionary.

(57) 要約: 単語の省略的な言い換え表現に対しても高い認識率で認識することが可能な音声認識用辞書を効率的に作成する音声認識用辞書作成装置(10)であって、1以上の単語から構成される認識対象語を構成単語に分割する単語分割部(2)、分割された構成単語ごとの読みに基づいて、構成単語ごとのモーラ列を生成するモーラ列取得部(3)、モーラを用いた省略語の生成規則を格納している省略語生成規則格

[続葉有]

WO 2004/044887 A1



(81) 指定国(国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:  
— 國際調査報告書

(84) 指定国(広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

## 明細書

## 音声認識用辞書作成装置および音声認識装置

## 技術分野

5 この発明は、不特定話者を対象とした音声認識装置に用いられる辞書を作成する音声認識用辞書作成装置およびその辞書を用いて音声を認識する音声認識装置等に関する。

## 背景技術

10 従来より、不特定話者を対象とした音声認識装置においては、認識語彙を規定する音声認識用辞書が不可欠である。認識対象語彙がシステム設計時において規定可能な場合には、事前に作成した音声認識用辞書を用いるが、語彙が規定できない場合、あるいは動的に変更されるべきである場合においては、人手による入力、または自動的に文字列情報から  
15 音声認識用語彙を作成し辞書に登録する。例えばテレビ番組切替装置における音声認識装置では、番組情報を含んだ文字列情報の形態素解析を行ってその表記の読みを求め、得られた読みを上記音声認識用辞書に登録する。例えば「NHKニュース10」という番組に対して、その読みである「えぬえいちけいにゅーすてん」を当該番組を表わす単語として  
20 音声認識用辞書に登録する。これにより、「えぬえいちけいにゅーすてん」というユーザの発声に対して、チャンネルを「NHKニュース10」に切り替える機能を実現することが可能となる。

また、ユーザが完全な単語を発話しないことを考慮して、複合単語を構成する単語に分割し、これらを連接しなおした部分文字列からなる言  
25 い換え表現を辞書に登録するという方法がある（例えば、特開2002  
-41081号公報に開示された技術）。上記公報に記載された音声認識

用辞書作成装置では、文字列情報として入力した単語を解析し、総ての読み、総ての連接単語を考慮して発声単位／読みのペアを作成して音声認識用辞書に登録する。これにより、例えば上記「NHKニュース10」という番組名に対し、「えぬえいちけいにゅーす」、「にゅーすてん」という読みが辞書に登録されることとなり、ユーザによるこれらの発声を正しく処理されることが期待される。

さらに上記音声認識用辞書作成方法は、上記言い換え表現に付された読みの確からしさを示す尤度や、言い換え表現を構成する単語の出現順位、当該単語が言い換え表現中で利用される頻度などを考慮した重み付けを行って、音声認識用辞書に登録する方法を提示している。これによつて、言い換え表現としてより確からしい単語が音声照合によって選択されることを期待する。

このように上記従来の音声認識用辞書作成方法は、入力された文字列情報を解析してあらゆる組合せの単語列を再構築し、これを当該単語の言い換え表現としてその読みを音声認識用辞書に登録することによって、正式な単語の発声だけでなく、ユーザによる任意の省略的な発声にも対処できることを目指すものである。

しかしながら、上記従来の音声認識用辞書作成方法は、次に挙げるような問題を有する。

すなわち、まず第1に、全網羅的にあらゆる組み合わせの文字列を生成した場合、その数は膨大なものとなる。このため、それら総てを音声認識用辞書に登録した場合、辞書が巨大となり、計算量の増加と音韻的に類似した多数の単語の登録により認識率の低下を招く恐れがある。さらに、異なる単語から生成された上記言い換え表現が同じ文字列・同じ読みとなる可能性が高くなり、たとえこれらが正しく認識されようとも、ユーザの発声が本来いずれの単語を意図したものであったかを識別する

のは、極めて困難となる。

また上記従来の音声認識用辞書作成方法では、非常に多く登録される言い換え表現候補の中から、より尤もらしいものを選択する目的で、言い換え表現中に表われる単語に関わる尤度を主として用いて言い換え表現の重み付けを求めてている。しかし、例えば「金曜ドラマ」を省略して「きんどら」と発声するようなケースを考えた場合、言い換え表現を生成する尤度を決定する要因は、組み合わされて使われた単語以上に、使われた単語から抜き出された音韻の数や、各々の音韻の連接の日本語としての自然さによって影響されるものであるといったことが考慮されていない。このため、言い換え表現に対する尤度が適切な値とならないという問題がある。

さらに、単語の言い換え表現は、単語を特定した場合、ほぼ 1 対 1 に対応するものであり、とりわけユーザを限定した場合にはその傾向は極めて顕著になると考えられる。上記従来の音声認識用辞書作成方法では、こののような言い換え表現の使用履歴を考慮した言い換え表現生成の制御を行っていないため、生成され認識辞書に登録される言い換え表現の数を適切に抑えることができないという問題を有する。

### 発明の開示

そこで本発明は、単語の省略的な言い換え表現に対しても高い認識率で認識することが可能な音声認識用辞書を効率的に作成する音声認識用辞書作成装置、および、これによって作成された音声認識用辞書を用いた省リソースで高性能な音声認識装置を提供することを目的とする。

上記目的を達成するために本発明に係る音声認識用辞書作成装置は、音声認識用辞書を作成する音声認識用辞書作成装置であって、1 以上の単語から構成される認識対象語について、発声のし易さを考慮したル

ルに基づいて、前記認識対象語の省略語を生成する省略語生成手段と、生成された省略語を前記認識対象語とともに前記音声認識用辞書として記憶する語彙記憶手段とを備えることを特徴とする。これによって、発声のし易さを考慮したルールに基づいて、前記認識対象語の省略語が生成され音声認識辞書として登録されるので、単語の省略的な言い換え表現に対しても高い認識率で認識することが可能な音声認識用辞書を効率的に作成する音声認識用辞書作成装置が実現される。

ここで、前記音声認識用辞書作成装置はさらに、前記認識対象語を構成単語に分割する単語分割手段と、分割された構成単語ごとの読みに基づいて、構成単語ごとのモーラ列を生成するモーラ列生成手段とを備え、前記省略語生成手段は、前記モーラ列生成手段によって生成された構成単語ごとのモーラ列に基づいて、構成単語ごとのモーラ列からモーラを取り出して連接することにより、1個以上のモーラからなる省略語を生成してもよい。このとき、前記省略語生成手段は、モーラを用いた省略語の生成規則を格納している省略語生成規則格納部と、前記構成単語ごとのモーラ列からモーラを取り出して連接することにより、1個以上のモーラからなる省略語の候補を生成する候補生成部と、生成された省略語の候補に対して、前記省略語生成規則格納部に格納された生成規則を適用することで、最終的に生成する省略語を決定する省略語決定部とを有してもよい。

上記構成によれば、構成単語のモーラ列から部分モーラ列を抽出し、これらを連接して省略語表現を構築する規則を事前に構築しておくことによって、新たな認識対象語に対しても可能性の高い省略語表現を生成することを可能とし、これを認識語彙として認識用辞書に登録することによって、認識対象語だけでなく当該単語の省略語表現の発声に対しても正しく認識できる音声認識装置を実現することが可能な音声認識用辞

書作成装置が作成される。

また、前記省略語生成規則格納部には、複数の生成規則が格納され、前記省略語決定部は、生成された省略語の候補について、前記省略語生成規則格納部に格納された複数の規則それぞれに対する尤度を算出し、  
5 算出した尤度を総合的に勘案することによって発声確率を決定し、前記語彙記憶手段は、前記省略語決定部によって決定された省略語および発声確率を前記認識対象語とともに記憶してもよい。ここで、前記省略語決定部は、前記複数の規則それぞれに対する尤度に、対応する重み付け係数を乗じて得られる値を合計することによって前記発声確率を決定し  
10 てもよい。そして、前記省略語決定部は、前記省略語の候補に対する発声確率が一定のしきい値を超える場合に、最終的に生成する省略語と決定してもよい。

上記構成によれば、認識対象語に対して生成される1語以上の省略語について各々発声確率が計算され、上記音声認識用辞書に省略語と関連付けられて格納される。これによって、1語の認識対象語に対して2語以上の省略語が生成された場合でも、それらから1語のみを絞り込むことなく、計算された発声確率に応じた重みを夫々の省略語に与えることが可能となり、比較的省略語として使われにくく予想される省略語に対しては低い確率が与えられ、音声との照合において高い認識精度を呈  
20 することのできる音声認識装置を実現できる音声認識用辞書を作成することができる。

また、前記省略語生成規則格納部には、単語の係り受けに関する第1の規則が格納され、前記省略語決定部は、前記第1の規則に基づいて、前記候補の中から最終的に生成する省略語を決定してもよい。例えば、  
25 前記第1の規則には、修飾語と被修飾語とを対にしてすることによって省略語を生成するという条件が含まれてもよいし、省略語を構成する修飾語

と被修飾語との距離と前記尤度との関係が含まれてもよい。

上記構成によれば、認識対象語に対応する省略語を生成する際に、認識対象語を構成する単語間の関係を考慮することが可能となり、構成単語間の関係に基づいた省略語を生成することが可能となる。これにより、  
5 認識対象語に含まれる構成単語中で、省略語に含まれる可能性の低い単語を除外したり、逆に省略語に含まれる可能性の高い単語を重点的に用いたりすることが可能となって、より適切な省略語を生成することができ、使用の可能性の低い省略語を認識用辞書に登録することを避け、高い認識精度を有する音声認識装置を実現できる音声認識用辞書を作成す  
10 ることができる。

また、前記省略語生成規則格納部には、省略語を生成するときに構成単語のモーラ列から取り出される部分モーラ列の長さおよび構成単語における位置の少なくとも1つに関する第2の規則が格納され、前記省略語決定部は、前記第2の規則に基づいて、前記候補の中から最終的に生  
15 成する省略語を決定してもよい。たとえば、前記第2の規則には、前記部分モーラ列の長さを示すモーラ数と前記尤度との関係が含まれてもよいし、前記部分モーラ列の構成単語における位置を示す構成単語の先頭からの距離に対応するモーラ数と前記尤度との関係が含まれてもよい。

上記構成によれば、当該単語を構成する単語の部分モーラを連接して  
20 省略語を生成する際の、抜き出した部分モーラ列の数や、各モーラの出現位置、生成された省略語の総モーラ数を考慮することが可能となる。これにより、複数の単語から構成される単語や長い単語を音韻的に短く切り詰めて省略語を生成する際の音韻の抽出に関わる一般的な傾向を、モーラという日本語等の言語における音韻のリズムの基本単位を用いて  
25 規則化することが可能となる。このため、認識対象語に対する省略語を生成する場合において、より適切な省略語を生成することができ、使用

の可能性の低い省略語を認識用辞書に登録することを避け、高い認識精度を有する音声認識装置を実現できる音声認識用辞書を作成することができる。

また、前記省略語生成規則格納部には、省略語を構成する部分モーラ  
5 列の連なりに関する第3の規則が格納され、前記省略語決定部は、前記  
第3の規則に基づいて、前記候補の中から最終的に生成する省略語を決  
定してもよい。たとえば、前記第3の規則には、連接された2つの部分  
モーラ列における前に位置する部分モーラ列の最後のモーラと後に位置  
する部分モーラ列の先頭のモーラとの組み合わせと前記尤度との関係が  
10 含まれてもよい。

上記構成によれば、複数の単語からなる単語や長い単語から省略語を  
生成する際に、音韻列が日本語等の言語として自然であるものが好まれ  
るという一般的な傾向を、モーラの連接確率という形で規則化するこ  
とが可能となる。これにより、認識対象語から省略語を生成する場合にお  
15 いて、より適切な省略語を生成することができ、使用の可能性の低い省  
略語を認識用辞書に登録することを避け、高い認識精度を有する音声認  
識装置を実現できる音声認識用辞書を作成することができる。

また、前記音声認識用辞書作成装置は、さらに、認識対象語を含んだ  
文字列情報から認識対象語を抽出する条件を格納している抽出条件格納  
20 手段と、認識対象語を含んだ文字列情報を取得する文字列情報取得手段  
と、前記抽出条件格納手段に格納されている条件に従って、前記文字列  
情報取得手段によって取得された文字列情報から認識対象語を抽出し、  
前記単語分割手段に送出する認識対象語抽出手段とを備えてよい。

上記構成によれば、文字列情報中から認識対象語を抽出する条件に応  
25 じて、適切に認識対象語を抽出し、かつ当該単語に対応する省略語を自  
動的に作成して、音声認識用辞書に格納することが可能となる。さらに

作成された各省略語について、省略語の生成に適用された規則に応じた尤度を基にした発声確率が計算され、この発声確率も同時に音声認識用辞書に格納される。これによって、文字列情報から自動的に作成された  
1語以上の省略語に対して、各々発声確率が与えらることとなり、音声  
5との照合において高い認識精度を呈することのできる音声認識装置を実現できる音声認識用辞書を作成することができる。

また、上記目的を達成するために、本発明に係る音声認識装置は、入力された音声を、音声認識用辞書に登録されている語彙に対応するモデルによって照合を行って認識する音声認識装置であって、前記音声認識  
10用辞書作成装置によって作成された音声認識用辞書を用いて前記音声を認識することを特徴とする。

上記構成によれば、事前に構築された音声認識用辞書内の語彙だけでなく、本発明に係る音声認識用辞書作成装置によって作成された、文字列情報から抽出された認識対象語およびこれから生成された省略語が格  
15納された音声認識用辞書内の語彙も認識の照合の対象とすることが可能となる。これによって、命令語のような固定的な語彙に加えて、検索キーワードのように文字列情報から抽出されるべき語彙、およびその省略語のいずれの語彙が発声された場合においても、正しく認識される音声認識装置を実現することが可能となる。

20 ここで、本発明に係る音声認識装置は、入力された音声を、音声認識用辞書に登録されている語彙に対応するモデルによって照合を行って認識する音声認識装置であって、前記音声認識用辞書作成装置を備え、前記音声認識用辞書作成装置によって作成された音声認識用辞書を用いて前記音声を認識してもよい。

25 上記構成によれば、搭載されている音声認識用辞書作成装置に文字列情報を入力することによって、自動的に認識対象語を抽出、およびその

省略語を生成して、音声認識用辞書に格納する。音声認識用辞書に格納されたこれら語彙は、音声認識装置において音声との照合を行うことが可能となるため、可変的に追加・変更するべき語彙を持つ音声認識装置において、その語彙およびその省略語を、文字列情報中から自動的に取得し、音声認識用辞書に登録することを可能とする。

ここで、前記音声認識用辞書には、前記省略語と当該省略語の発声確率とが前記認識対象語とともに登録され、前記音声認識装置は、前記音声認識用辞書に登録されている発声確率を考慮して前記音声の認識を行ってもよい。そして、前記音声認識装置は、前記音声の認識結果である候補とともに当該候補の尤度を生成し、生成した尤度に前記発声確率に対応する尤度を加算し、得られた加算値に基づいて前記候補を最終的な認識結果として出力してもよい。

上記構成によれば、文字列情報中から認識対象語を抽出しあつその省略語を生成する過程で、各省略語の発声確率も計算されて音声認識用辞書に格納される。音声認識装置では、音声の照合の際に各省略語の発声確率を考慮した照合を行うことが可能となり、省略語として比較的可能性の低いものについては、低めの確率が与えられるといった制御が可能となり、不自然な省略語の湧き出しによる音声認識の正解確率の低下を抑えることができる。

また、前記音声認識装置は、さらに、前記音声に対して認識した省略語と当該省略語に対応する認識対象語とを使用履歴情報として格納する省略語使用履歴格納手段と、前記省略語使用履歴格納手段に格納された使用履歴情報に基づいて、前記省略語生成手段による省略語の生成を制御する省略語生成制御手段を備えてよい。たとえば、前記音声認識用辞書作成装置の省略語生成手段は、モーラを用いた省略語の生成規則を格納している省略語生成規則格納部と、前記構成単語ごとのモーラ列か

らモーラを取り出して連接することにより、1個以上のモーラからなる省略語の候補を生成する候補生成部と、生成された省略語の候補に対して、前記省略語生成規則格納部に格納された生成規則を適用することで、最終的に生成する省略語を決定する省略語決定部とを有し、前記省略語生成制御手段は、前記省略語生成規則格納部に格納される生成規則を変更、削除または追加することによって前記省略語の生成を制御してもよい。

同様に、前記音声認識装置は、さらに、前記音声に対して認識した省略語と当該省略語に対応する認識対象語とを使用履歴情報として格納する省略語使用履歴格納手段と、前記省略語使用履歴格納手段に格納された使用履歴情報に基づいて、前記音声認識用辞書に格納されている省略語に対する編集を行う辞書編集手段とを備えててもよい。たとえば、前記音声認識用辞書には、前記省略語と当該省略語の発声確率とが前記認識対象語とともに登録され、前記辞書更新手段は、前記省略語の発声確率を変更することによって前記省略語に対する編集を行ってもよい。

上記構成によれば、ユーザの過去の省略語の使用に関する履歴情報を元に、ユーザの省略語使用に関する傾向を考慮して上記省略語生成規則を制御することが可能となる。これは、ユーザの省略語利用には一定の傾向があり、また、同一の単語に対しては多くても2語程度の省略語しか用いることはないということに着目したものである。すなわち、省略語新規生成においては、過去の省略語利用から利用傾向の強い省略語だけを生成することが可能となる。また、すでに上記認識用辞書に記憶された省略語についても、同一の単語から複数の省略語が生成された場合において、ある省略語のみが利用され、その他の省略語が利用されないことが明らかとなれば、辞書からこれらを削除することが可能となる。このような機能により、過剰な省略語が、上記認識用辞書に登録される

のを防ぎ、音声認識の性能の低下を抑えることが可能となる。また、異なる認識対象語に対して生成されたそれぞれの省略語の中に、共通の省略語が存在するようなケースにおいても、過去のユーザの具体的な省略語の使用情報から、いずれの認識対象語を意図したものであるかを予測  
5 することが可能となる。

なお、本発明は、上記のような音声認識用辞書作成および音声認識装置として実現することができるだけでなく、これらの装置が備える特徴的な手段をステップとする音声認識用辞書作成方法および音声認識方法として実現したり、それらのステップをコンピュータに実行させるプログラムとして実現したりすることができる。そして、そのようなプログラムは、CD-ROM等の記録媒体やインターネット等の通信媒体を介して配布することができるのは言うまでもない。  
10

#### 図面の簡単な説明

15 図1は、本発明の実施の形態1における音声認識用辞書作成装置の構成を示す機能ブロック図である。

図2は、同音声認識用辞書作成装置による辞書作成処理を示すフローチャートである。

20 図3は、図2に示された省略語生成処理(S23)の詳細な手順を示すフローチャートである。

図4は、同音声認識用辞書作成装置の省略語生成部が有する処理テーブル(一時的に発生する中間データ等を記憶するテーブル)を示す図である。

25 図5は、同音声認識用辞書作成装置の省略語生成規則格納部に格納されている省略語生成規則の例を示す図である。

図6は、同音声認識用辞書作成装置の語彙記憶部に格納されている音

声認識用辞書の例を示す図である。

図 7 は、本発明の実施の形態 2 における音声認識装置の構成を示す機能ブロック図である。

図 8 は、同音声認識装置の学習機能を示すフローチャートである。

5 図 9 は、同音声認識装置の応用例を示す図である。

図 10 (a) は、中国語の認識対象語から音声認識用辞書作成装置 10 によって生成される省略語の例を示す図であり、図 10 (b) は、英語の認識対象語から音声認識用辞書作成装置 10 によって生成される省略語の例を示す図である。

10

#### 発明を実施するための最良の形態

以下、本発明の実施の形態について、図面を参照しながら説明する。

##### (実施の形態 1)

図 1 は、実施の形態 1 における音声認識用辞書作成装置 10 の構成を示す機能ブロック図である。この音声認識用辞書作成装置 10 は、認識対象語からその省略語を生成し、辞書として登録する装置であり、プログラムや論理回路として実現される認識対象語解析部 1 および省略語生成部 7 と、ハードディスクや不揮発性メモリ等の記憶装置等によって実現される解析用単語辞書格納部 4、解析規則格納部 5、省略語生成規則格納部 6 および語彙記憶部 8 から構成される。

解析用単語辞書格納部 4 は、認識対象語を構成単語に分割するための単位単語（形態素）およびその音韻系列の定義（音韻情報）に関する辞書を予め格納している。解析規則格納部 5 は、認識対象語を解析用単語辞書格納部 4 に格納されている単位単語に分割するための規則（構文解析用の規則）を予め格納している。

省略語生成規則格納部 6 は、事前に構築された単語の省略語を生成す

るための複数の規則、つまり、発声のし易さを考慮した複数の規則を予め格納している。これらの規則の中には、例えば、認識対象語を構成する単語そのものや、これらの係り受け関係を元に、構成単語中から部分モーラ列を抽出する単語を決定する規則や、構成単語から抽出する部分モーラの抽出位置や、抽出数、ならびにそれらを組み合わせた際の総モーラ数を元に、適切な部分モーラの抽出を行う規則、さらに、抽出したモーラを連接した際のモーラ連接の自然さを元に、部分モーラの連接を行う規則などが含まれる。  
5

なお、「モーラ」とは、1音(1拍)と考えられている音韻のことであり、  
10 日本語であれば、ひらかな表記した時のひらかな1文字1文字に概ね対応する。また、俳句の 5・7・5 をカウントする時の1音に対応する。ただし、拗音(小さいや ゆ ょ の付く音)、促音(小さい っ / つまつた音)、撥音(ん)については、1音(1拍)として発音されるか否かによって、独立した1つのモーラとして取り扱われたり、そうでなかったりする。  
15 例えば、「東京」であれば、4つのモーラ「と」、「う」、「きょ」、「う」から構成され、「札幌」であれば、4つのモーラ「さ」、「っ」、「ぼ」、「ろ」から構成され、「群馬」であれば、3つのモーラ「ぐ」、「ん」、「ま」から構成される。

認識対象語解析部1は、この音声認識用辞書作成装置10に入力された認識対象語に対して形態素解析・構文解析・モーラ解析等を行う処理部であり、単語分割部2とモーラ列取得部3とから構成される。単語分割部2は、解析用単語辞書格納部4に格納された単語の情報および解析規則格納部5に格納された構文解析規則に従って、入力された認識対象語を、その認識対象語を構成する単語(構成単語)に分割するとともに、  
20 分割した構成単語の係り受け関係(修飾語と被修飾語の関係を示す情報)も生成する。モーラ列取得部3は、解析用単語辞書格納部4に格納され  
25

た単語の音韻情報に基づいて、単語分割部2で生成された構成単語ごとに、モーラ列を生成する。この認識対象語解析部1による解析結果、つまり、単語分割部2から生成される情報（認識対象語を構成する単語の情報と単語間の係り受け関係）およびモーラ列取得部3から生成される  
5 情報（各構成単語の音韻系列を表わすモーラ列）は省略語生成部7に送られる。

省略語生成部7は、省略語生成規則格納部6に格納された省略語生成規則を用いて、認識対象語解析部1から送られてきた認識対象語に関する情報から、その認識対象語の省略語を0語以上生成する。具体的には、  
10 認識対象語解析部1から送られてきた各単語のモーラ列を係り受け関係に基づいて組み合わせたりすることで、省略語の候補を生成し、生成した省略語の候補それぞれについて、省略語生成規則格納部6に格納された規則ごとの尤度を算出する。そして、一定の重み付けを乗じたうえで、各尤度を合計することによって、候補ごとの発声確率を計算し、一定以上の発声確率をもつ候補を、最終的な省略語として、その発声確率および元の認識対象語と対応づけて語彙記憶部8に格納する。つまり、省略語生成部7によって一定以上の発声確率を持つと判断された省略語は、  
15 入力された認識対象語と同一の意味を持つ単語であることを示す情報、および、その発声確率とともに、音声認識用辞書として、語彙記憶部8に登録される。  
20

語彙記憶部8は、書き換え可能な音声認識用辞書を保持するとともに登録処理を行うものであり、省略語生成部7で生成された省略語および発声確率を、この音声認識用辞書作成装置10に入力された認識対象語と対応づけたうえで、それら認識対象語、省略語および発声確率を音声  
25 認識用辞書として登録する。

次に、以上のように構成された音声認識用辞書作成装置10の動作に

ついて、具体例とともに説明する。

図2は、音声認識用辞書作成装置10の各部によって実行される辞書作成処理動作のフローチャートである。なお、本図における矢印の左側には、認識対象語として「朝の連續ドラマ」が入力された場合の具体的な中間データや最終データ等が示され、右側には参照または格納の対象となるデータ名が記されている。  
5

まず、ステップS21において、認識対象語が認識対象語解析部1の単語分割部2に読み込まれる。単語分割部2は、その認識対象語を、解析用単語辞書格納部4に格納された単語の情報と、解析規則格納部5に10格納された単語分割規則に従って、構成単語に分割するとともに、各構成単語の係り受け関係を求める。つまり、形態素解析と構文解析を行う。これによって認識対象語「朝の連續ドラマ」は、例えば、「朝」、「の」、「連續」、「ドラマ」という構成単語に分割され、その係り受け関係として、(朝) → ((連續) → (ドラマ)) という関係が生成される。なお、15この係り受け関係の表記において、矢印の元が修飾語を、矢印の先が被修飾語を示している。

ステップS22では、モーラ列取得部3は、単語分割処理ステップS21において分割された各構成単語に対して、その音韻系列としてのモーラ列を付与する。このステップでは、構成単語の音韻系列を得るために、解析用単語辞書格納部4に格納された単語の音韻情報が利用される。その結果、単語分割部2で得られた構成単語「朝」、「の」、「連續」、「ドラマ」に対して、それぞれ、「アサ」、「ノ」、「レンゾク」、「ドラマ」というモーラ列が付与される。このようにして得られたモーラ列は、上記ステップS21で得られた構成単語および係り受け関係の情報と共に、省25略語生成部7に送出される。

ステップS23では、認識対象語解析部1から送られてくる構成単語、

係り受け関係およびモーラ列から、省略語生成部 7 は、省略語を生成する。ここでは、省略語生成規則格納部 6 に格納された 1 つ以上の規則が適用される。これらの規則の中には、認識対象語を構成する単語そのものや、これらの係り受け関係を元に、構成単語中から部分モーラ列を抽出する単語を決定する規則や、構成単語から抽出する部分モーラの抽出位置や、抽出数、ならびにそれらを組み合わせた際の総モーラ数を元に、適切な部分モーラの抽出を行う規則、さらに、抽出したモーラを連接した際のモーラ連接の自然さを元に、部分モーラの連接を行う規則などが含まれている。省略語生成部 7 は、省略語の生成に適用される規則ごとに、規則の一致度を示す尤度を計算し、複数の規則で計算した尤度を総合することによって、生成した省略語の発声確率を計算する。その結果、例えば、省略語として、「アサドラ」、「レンドラ」、「アサレンドラ」が生成され、この順に高い発声確率が与えられる。

ステップ S 2 4 では、語彙記憶部 8 は、省略語生成部 7 が生成した省略語および発声確率の組を認識対象語と対応づけて音声認識用辞書に格納する。このようにして、認識対象語の省略語とその発声確率が格納された音声認識用辞書が作成される。

次に、図 2 に示された省略語生成処理 (S 2 3) の詳細な手順を図 3 ~ 図 5 を用いて説明する。図 3 は、その詳細な手順を示すフローチャートであり、図 4 は、省略語生成部 7 が有する処理テーブル（一時的に発生する中間データ等を記憶するテーブル）を示し、図 5 は、省略語生成規則格納部 6 に格納されている省略語生成規則 6 a の例を示す図である。

まず、省略語生成部 7 は、認識対象語解析部 1 から送られてくる構成単語、係り受け関係およびモーラ列に基づいて、省略語の候補を生成する（図 3 の S 3 0）。具体的には、認識対象語解析部 1 から送られてきた構成単語の係り受け関係が示す修飾語と被修飾語からなる総ての組み合

わせを省略語の候補として生成する。このとき、図 4 の処理テーブルにおける「省略語の候補」に示されるように、修飾語および被修飾語それぞれについて、構成単語のモーラ列だけでなく、その一部を欠落させた部分モーラ列も用いられる。例えば、修飾語「レンゾク」と被修飾語「ドラマ」との組み合わせについては、「レンゾクドラマ」だけでなく、「レンゾクドラ」、「レンドラマ」、「レンドラ」等の 1 個以上のモーラを欠落させてできる総てのモーラ列が省略語の候補として生成される。  
5

次に、省略語生成部 7 は、生成した省略語の候補それぞれについて(図 3 の S 3 1 ~)、省略語生成規則格納部 6 に格納されている省略語生成規 10 則ごとの尤度を算出し(図 3 の S 3 2 ~ S 3 4)、一定の重み付けの下で各尤度を合計することによって発声確率を算出する(図 3 の S 3 5)と 15 いう処理を繰り返す(図 3 の S 3 0 ~ S 3 6)。

例えば、省略語生成規則の 1 つとして、図 5 のルール 1 に示されるよ 15 うに、係り受け関係に関する規則であって、修飾語と被修飾語とをこの順で結合すること、および、修飾語と被修飾語との距離(図 4 の上部に示される係り受け関係図における段数)が小さいほど高い尤度を示す関数等が定義されているとする。すると、省略語生成部 7 は、各候補省略語について、このようなルール 1 に対応する尤度を算出する。例えば、 20 「レンドラ」について、修飾語と被修飾語がこの順で結合された省略語であることを確認したうえで(そうでなければ、尤度を 0 とする)、修飾語「レン」と被修飾語「ドラ」との距離(ここでは、「レン(ゾク)」が「ドラ(マ)」を修飾しているので 1 段)を特定し、その距離に対応する尤度(ここでは、0.102)を上述の関数に従って特定する。

なお、「アサドラ」であれば、修飾語「アサ」と被修飾語「ドラ」との 25 距離は、「アサ」が「レンゾクドラマ」を修飾していることから、2 段となり、また、「アサレンドラ」であれば、修飾語と被修飾語との距離は、

上記「レンドラ」と「アサドラ」の両方の係り受け関係を有することから、それら 2 つの距離の平均値、つまり、1.5 段となる。

また、省略語生成規則の他の例として、図 5 のルール 2 に示されるように、部分モーラ列に関する規則であって、部分モーラ列の位置に関するルールと長さに関するルール等が定義されているとする。具体的には、部分モーラ列の位置に関するルールとして、修飾語または被修飾語として採用されたモーラ列（部分モーラ列）が元の構成単語の先頭に近い位置であるほど高い尤度を示すというルール、つまり、先頭からの距離（元の構成単語の先頭と部分モーラ列の先頭に挟まれたモーラ数） vs 尤度の関係を示す関数等が定義されている。また、部分モーラ列の長さに関するルールとして、部分モーラ列を構成するモーラの数が 2 に近いほど高い尤度を示すというルール、つまり、部分モーラ列の長さ（モーラ数） vs 尤度の関係を示す関数が定義されている。省略語生成部 7 は、各候補省略語について、このようなルール 2 に対応する尤度を算出する。例えれば、「アサドラ」について、部分モーラ列「アサ」および「ドラ」それぞれについて、構成単語「アサ」および「ドラマ」における位置および長さを特定し、上述の関数に従って各尤度を特定し、それら尤度の平均値をルール 2 に対する尤度（ここでは、0.128）とする。

また、省略語生成規則の他の例として、図 5 のルール 3 に示されるように、音韻の連なりに関する規則であって、部分モーラ列の結合部分に関するルール等が定義されているとする。ここで、部分モーラ列の結合部分に関するルールとして、結合されている 2 つの部分モーラ列における前の部分モーラ列の最後尾のモーラと後の部分モーラ列の先頭のモーラとの結合が、不自然な音韻の組み合わせ（発音しにくい音韻）である場合に低い尤度となるようなデータテーブルが定義されている。省略語生成部 7 は、各候補省略語について、このようなルール 3 に対応する尤

度を算出する。具体的には、各部分モーラ列の結合部分がルール 3 に登録された不自然な連なりのいずれかに属するか否かを判断し、属する場合には、その連なりに対応する尤度を割り当て、そうでない場合には、デフォルト値の尤度（ここでは、0.050）を割り当てる。例えば、  
5 「アサレンドラ」について、部分モーラ列「アサ」と「レン」との結合部分「サレ」がルール 3 に登録された不自然な連なりに属するか否かを判断する。ここでは、いずれにも属さないので、尤度をデフォルト値（0.050）とする。

このようにして、省略語の候補それぞれについて省略語生成規則ごとの尤度を算出すると、省略語生成部 7 は、図 3 のステップ S 3.5 に示される発声確率  $P(w)$  の算出式に従って、各尤度  $\times$  に重み付け（図 5 に示された対応するルールごとの重み  $\alpha$ ）を乗じて合計することによって、候補ごとの発声確率を算出する（図 3 の S 3.5）。

最後に、省略語生成部 7 は、総ての候補の中から、予め設定された一定のしきい値を超える発声確率を持つものを特定し、それらを最終的な省略語として発声確率とともに語彙記憶部 8 に出力する（図 3 の S 3.7）。これによって、語彙記憶部 8 において、図 6 に示されるように、認識対象語の省略語と発声確率とが含まれる音声認識用辞書 8a が作成される。

以上のようにして作成された音声認識用辞書 8a は、認識対象語だけでなく、その省略語が発声確率とともに登録されている。したがって、この音声認識用辞書作成装置 10 によって作成された音声認識用辞書を用いることで、正式な単語を発声した場合においても、その省略語を発声した場合においても、同じ意図の発声であることを検出し、高い認識率で音声を認識することが可能な音声認識装置が実現される。例えば、  
20 上記「朝の連続ドラマ」の例では、ユーザが「アサノレンゾクドラマ」と発声した場合でも、「アサドラ」と発声した場合でも「朝の連続ドラマ」  
25

と認識し、同様に機能することができる音声認識装置のための音声認識用辞書が作成される。

(実施の形態 2)

実施の形態 2 は、実施の形態 1 における音声認識用辞書作成装置 10 を搭載し、この音声認識用辞書作成装置 10 によって作成された音声認識用辞書 8a を用いる音声認識装置の例に関する。本実施の形態では、認識対象語を文字列情報から自動的に抽出し、これを音声認識用辞書に格納する辞書更新機能を有し、かつ、ユーザによる過去の省略語使用の履歴に基づいた情報を用いて省略語の生成を制御することで、利用する可能性の低い省略語が認識用辞書に登録されるのを抑える機能を有する音声認識装置に関する。なお、文字列情報とは、音声認識装置による認識の対象となる語（認識対象語）を含む情報であり、例えば、デジタルTV放送を視聴する視聴者が発した番組名に基づく番組の自動切替を行う音声認識装置の応用例であれば、番組名が認識対象語となり、放送局から放送されてくる電子番組データが文字列情報となる。

図 7 は、実施の形態 2 における音声認識装置 30 の構成を示す機能ブロック図である。この音声認識装置 30 は、実施の形態 1 における音声認識用辞書作成装置 10 に加えて、文字列情報取込部 17、認識対象語抽出条件格納部 18、認識対象語抽出部 19、音声認識部 20、ユーザ I/F 部 25、省略語使用履歴格納部 26 および省略語生成規則制御部 27 から構成される。なお、音声認識用辞書作成装置 10 は、実施の形態 1 のものと同一であり、その説明を省略する。

文字列情報取込部 17、認識対象語抽出条件格納部 18、認識対象語抽出部 19 は、認識対象語が含まれる文字列情報から認識対象語を抽出するためのものである。この構成によれば、文字列情報取込部 17 は、認識対象語が含まれた文字列情報を取り込み、続く認識対象語抽出部 1

9において、この文字列情報から認識対象語の抽出を行う。認識対象語を文字列情報から抽出するために、文字列情報は形態素解析された後に、認識対象語抽出条件格納部18に格納された認識対象語抽出条件に従つて抽出が行われる。抽出された認識対象語は、音声認識用辞書作成装置5 10に送出され、その省略語の作成と、認識辞書への登録が行われる。

これによって、本実施の形態の音声認識装置30では、電子番組データのような文字列情報から、番組名のような検索キーワードを自動的に抽出し、このキーワードおよびそこから生成された省略語のいずれを発声しても、正しく音声認識することのできる音声認識用辞書が作成され10 る。なお、認識対象語抽出条件格納部18に格納される認識対象語抽出条件とは、例えば、デジタル放送受信機に入力されるデジタル放送データ中の電子番組データを識別する情報や、電子番組データ中の番組名を識別する情報等である。

音声認識部20は、マイク等から入力された入力音声に対して、音声認識用辞書作成装置10で作成された音声認識用辞書に基づく音声認識を行う処理部であり、音響分析部21、音響モデル格納部22、固定語彙記憶部23、照合部24からなる。マイク等から入力された音声は、音響分析部21で周波数分析等が行われ、特徴パラメータの系列（メルケプストラム係数など）へと変換される。照合部24では、音響モデル20 格納部22に格納されたモデル（例えば、隠れマルコフモデルや混合ガウス分布モデルなど）を用いて、固定語彙記憶部23に格納された語彙（固定語彙）、または、語彙記憶部8に格納された語彙（通常語および省略語）を元に、各語彙を認識するためのモデルを合成しながら入力音声との合成を行う。その結果、高い尤度を得た単語が認識結果候補として25 ユーザI/F部25に送出される。

このような構成により、この音声認識部20により、機器制御コマン

ド（例えば、番組切替における発声「切り替え」）等のシステム構築時に決定可能な語彙を固定語彙記憶部23に格納しておき、番組切替のための番組名のように番組名の変化に応じて可変的に変更する必要のある語彙を語彙記憶部8に格納しておくことで、双方の語彙を同時に認識する  
5 ことが可能となる。

また、語彙記憶部8には、省略語だけでなく、発声確率も格納されている。この発声確率は、照合部24において音声の照合を行う際に利用され、発声確率の低い省略語は認識されにくくすることによって、省略語の過剰な湧き出しによる音声認識装置の性能の低下を抑えることが可能となっている。  
10 例えば、照合部24は、入力された音声と語彙記憶部8に格納された語彙との相関を示す尤度に、語彙記憶部8に格納された発声確率に対応する尤度（たとえば、発声確率の対数値）を加算し、得られた加算値を認識結果に対する最終的な尤度とし、その最終的な尤度が一定のしきい値を超える場合に、その語彙を認識結果候補としてユ  
15 ザI/F部25に送出する。なお、一定のしきい値を超える認識結果候補が複数ある場合には、それらのうち、尤度が最も大きいものから一定順位内のものだけをユーザI/F部25に送出する。

ところで、このような音声認識用辞書作成装置10によっても、複数の異なる認識対象語に対して、共通の音韻系列となる省略語が生成される可能性がある。これは、省略語生成規則に残るあいまい性のために生じる問題である。通常、ユーザは1つの省略語は1つの対応する認識対象語を意味する目的で利用していると考えられる。したがって、省略語生成規則に残るあいまい性を解消し、発声された省略語から適切な動作を提示できるとともに、長く使用することによって認識率が向上する学  
20 習機能を備えた音声認識装置が必要とされる。ユーザI/F部25、省略語使用履歴格納部26、省略語生成規則制御部27は、このような学  
25 習機能を備えた音声認識装置が必要とされる。

習機能のための構成要素である。

すなわち、ユーザ I／F 部 25 は、照合部 24 での音声照合の結果、認識結果候補を 1 つに絞り込むことができなかった場合、それら複数の候補をユーザに提示するとともに、ユーザから選択指示を取得する。例 5 えば、ユーザの発話に対して得られた複数の認識結果の候補（切替先となる複数の番組名）を TV 画面に表示する。ユーザは、リモコン等を用いて、その中から 1 つの正解候補を選択することで所望の動作（音声による番組の切り替え）を得ることができる。

このようにしてユーザ I／F 部 25 に送出された省略語、あるいは、10 ユーザ I／F 部 25 に送出された複数の省略語の中からユーザによって選択された省略語は、履歴情報として、省略語使用履歴格納部 26 に送出され格納される。省略語使用履歴格納部 26 に格納された履歴情報は、省略語生成規則制御部 27 において集計され、省略語生成規則格納部 6 に格納された省略語生成のための規則やパラメータ、また省略語の発声 15 確率を計算するためのパラメータを変更するために用いられる。同時に、ユーザの省略語使用によって、本来の単語とその省略語の間に 1 対 1 の対応関係が得られた場合には、その情報も省略語生成規則格納部に格納される。また、このような省略語生成規則格納部 6 の規則の追加・変更・削除についての情報は、語彙記憶部 8 にも送られ、既に登録済みの省略 20 語についての見直しが行われ、省略語の削除・変更が行われて、辞書の更新が行われる。

図 8 は、このような音声認識装置 30 の学習機能を示すフローチャートである。

ユーザ I／F 部 25 は、照合部 24 から送られてくる認識結果候補に、25 語彙記憶部 8 に格納された省略語が含まれている場合には、その省略語を省略語使用履歴格納部 26 に送ることで、省略語使用履歴格納部 26

に蓄積させる(S 4 0)。このとき、ユーザが選択した省略語については、その旨を示す情報を附加して省略語使用履歴格納部26に送る。

省略語生成規則制御部27は、一定期間が経過する度に、または、一定の情報量が省略語使用履歴格納部26に蓄積される度に、省略語使用  
5 履歴格納部26に蓄積された省略語を統計的に解析することで、規則性を生成する(S 4 1)。例えば、省略語の長さ(モーラ数)に関する頻度分布や省略語を構成するモーラの連なりに関する頻度分布等を生成する。また、ユーザの選択情報等に基づいて、例えば、番組名「朝の連續ドラマ」を「レンドラ」と称していることが確認できた場合には、それら認  
10 識対象語と省略語との1対1の対応関係を示す情報も生成する。なお、このような規則性の生成を終えると、省略語生成規則制御部27は、省略語使用履歴格納部26の記憶内容を消去し、さらなる蓄積に備える。

そして、省略語生成規則制御部27は、生成した規則性に従って、省略語生成規則格納部6に格納されている省略語生成規則を追加、変更または削除をする(S 4 2)。例えば、省略語の長さに関する頻度分布に基づいて、図5のルール2に含まれる部分モーラ列の長さに関する規則(分布を示す関数のパラメータのうち、平均値を特定するパラメータ等)を修正する。また、認識対象語と省略語との1対1の対応関係を示す情報が生成された場合には、その対応関係を新たな省略語生成規則として登  
20 録する。

省略語生成部7は、このように追加・変更・削除された省略語生成規則に従って、認識対象語に対する省略語の生成を繰り返すことで、語彙記憶部8に格納されている音声認識用辞書の見直しを行う(S 4 3)。例えば、新たな省略語生成規則に従って省略語「アサドラ」の発声確率を計算し直した場合には、その発声確率を更新したり、認識対象語「朝の連續ドラマ」に対してユーザが省略語として「レンドラ」を選択した場

合には、省略語「レンドラ」の発声確率を増加させたりする。

このようにして、本音声認識装置30により、省略語を含めた音声認識が行われるだけでなく、認識結果に従って省略語生成規則が更新され、音声認識用辞書が改定されていくので、使用時間とともに認識率が向上するという学習機能が発揮される。  
5

図9(a)は、このような音声認識装置30の応用例を示す図である。

ここでは、音声によるTV番組の自動切替システムが示されている。このシステムは、音声認識装置30が内蔵されたSTB(Set Top Box; デジタル放送受信機)40と、TV受像機41と、無線マイクの機能を備えるリモコン42とから構成される。ユーザの発話は、リモコン42のマイクを介して音声データとしてSTB40に送信され、STB40に内蔵された音声認識装置30によって音声認識され、その認識結果に応じて、番組切替が行われる。  
10  
15

例えば、ユーザが、「レンドラニキリカエ」と発話したとする。すると、その音声はリモコン42を介して、STB40に内蔵された音声認識装置30に送信される。音声認識装置30の音声認識部20は、図9(b)の処理手順に示されるように、入力された音声「レンドラニキリカエ」に対して、語彙記憶部8および固定語彙記憶部23を参照することで、可変語彙「レンドラ」(つまり、認識対象語「朝の連続ドラマ」)および固定語彙「キリカエ」が含まれていることを検出する。その結果に基づいて、STB40は、予め放送データとして受信し保持している電子番組データの中に、現在放送中の番組「朝の連続ドラマ」が存在することを確認したうえで、その番組(ここでは、チャネル6)を選局する切替制御を行う。  
20  
25

このように、本実施の形態の音声認識装置では、機器制御のための命令語のような固定的な語彙の認識と、番組検索のための番組名のような

可変的な語彙の認識が同時に見えるばかりでなく、固定語彙についても、可変語彙についても、さらにその省略語表現に対しても、機器の制御等と連動させることで、所望の処理を行うことができる。さらにユーザの過去の使用履歴を考慮した学習により、省略語生成過程のあいまい性を  
5 解消し、高い認識率を持つ音声認識用辞書を効率的に作成することが可能となる。

以上、本発明に係る音声認識用辞書作成装置および音声認識装置について、実施の形態に基づいて説明したが、本発明はこれらの実施の形態に限定されるものではない。

10 たとえば、実施の形態 1 及び 2 では、日本語を対象とした音声認識用辞書作成装置 10 及び音声認識装置 30 の例が示されたが、本発明は、日本語だけでなく、中国語や英語等の日本語以外の言語にも適用することができるのは言うまでもない。図 10 (a) は、中国語の認識対象語から音声認識用辞書作成装置 10 によって生成される省略語の例を示す図  
15 であり、図 10 (b) は、英語の認識対象語から音声認識用辞書作成装置 10 によって生成される省略語の例を示す図である。これらの省略語は、例えば、図 5 に示される省略語生成規則 6a、「認識対象語の先頭 1 シラブルを省略語とする」、「認識対象語を構成する各単語の先頭 1 シラブルを連結したものを省略語とする」等の省略語生成規則によって生成  
20 され得る。

また、実施の形態 1 の音声認識用辞書作成装置 10 は、発声確率の高い省略語を生成したが、省略されていない通常語についても生成対象としてもよい。たとえば、省略語生成部 7 は、省略語だけでなく、省略していない認識対象語に対応するモーラ列についても、予め定められた一定の発声確率とともに、語彙記憶部 8 の音声認識用辞書に固定的に登録してもよい。あるいは、音声認識装置において、その音声認識用辞書に

登録されている省略語だけでなく、音声認識用辞書のインデックスとなっている認識対象語についても認識対象に含めることで、省略語だけでなく、フルスペリングに対応する通常語についても同時に認識することが可能となる。

5 また、実施の形態 1において、省略語生成規則制御部 27は、省略語生成規則格納部 6に格納された省略語生成規則の変更等を行ったが、直接、語彙記憶部 8の内容を変更してもよい。具体的には、語彙記憶部 8に格納されている音声認識用辞書 8aに登録されている省略語の追加、変更または削除をしたり、登録されている省略語の発声確率を増減させ  
10 てもよい。これによって、省略語使用履歴格納部 26に格納された使用履歴情報に基づいて、直接、音声認識用辞書が修正されることになる。

また、省略語生成規則格納部 6に格納される省略語生成規則および規則中の用語の定義としては、本実施の形態だけに限られない。たとえば、本実施の形態では、修飾語と被修飾語との距離は、係り受け関係図における段数を意味したが、このような定義に限られるものではなく、修飾語と被修飾語の意味的な継続性の良否を与えるような値を「修飾語と被修飾語との距離」と定義してもよい。例として、「(真っ赤な (夕陽))」と「(真っ青な (夕陽))」とでは、前者の方が意味的に自然であるので、前者の方が近い距離となるような尺度を採用してもよい。

20

また、実施の形態 2では、音声認識装置 30の適用例として、デジタル放送受信システムにおける自動番組切替が示されたが、このような自動番組切替は、放送システム等の一方向性の通信システムだけに限られず、インターネットや電話網等の双方向の通信システムにおける番組切替にも適用できるのは言うまでもない。たとえば、本発明に係る音声認識装置を携帯電話機に内蔵することで、ユーザが望むコンテンツの指

定を音声認識し、インターネット上のサイトからそのコンテンツをダウンロードするというコンテンツ配信システムを実現することができる。たとえば、ユーザが「クマピーラダウンロード」と発話すると、可変語彙「クマピー（「くまのピーさん」の省略語）」と固定語彙「ダウンロード」とが認識され、インターネット上のサイトから着メロ「くまのピーさん」が携帯電話機にダウンロードされる。

同様に、本発明に係る音声認識装置30は、放送システムやコンテンツ配信システム等の通信システムだけに限られず、スタンドアローンの機器にも適用することができる。たとえば、本発明に係る音声認識装置10 30をカーナビゲーション装置に内蔵させることで、運転者が発話した行先の地名等を音声認識し、その行先までの地図が自動表示されるという便利で安全性の高いカーナビゲーション装置が実現される。たとえば、運転しながら、「カドカドヒヨウジ」と発話すると、可変語彙「カドカド（「大阪府門真市大字門真」の省略語）」と固定語彙「ヒヨウジ」とが15 認識され、カーナビゲーションの画面に、「大阪府門真市大字門真」付近の地図が自動表示される。

以上のように、本発明によって、認識対象語の正式な発声だけでなくその省略語を発声した場合においても同様に動作する音声認識装置用の音声認識用辞書が作成される。また本発明では、日本語音声の発声リズムであるモーラに着目した省略語生成規則が適用され、さらにそれら省略語の発声確率を考慮した重み付けが付与されるので、無用な省略語の生成と認識辞書への登録を避けることが可能になるとともに、重み付けの併用によって、湧き出した省略語が音声認識装置の性能に悪影響を与えることが避けられる。

25 また、このような音声認識用辞書作成装置を搭載した音声認識装置では、省略語使用についてのユーザの履歴を音声認識用辞書作成部で利用

することで、省略語生成規則のあいまい性から生じる元単語対省略語の間の多対多の対応関係を解消することが可能となり、効率的な音声認識用辞書の構築が可能となる。

また、本発明に係る音声認識装置では、認識結果を音声認識用辞書の  
5 作成プロセスに反映するフィードバックが形成されているので、装置の使用に伴って認識率が向上していくという学習効果が発揮される。

このように、本発明によって、省略語を含む音声が高い認識率で認識されることとなり、放送番組の切替、携帯電話機に対する操作、カーナビゲーション装置に対する指示等が省略語を含む音声によって行われることとなり、本発明の実用的価値は極めて高い。  
10

#### 産業上の利用の可能性

本発明は、不特定話者を対象とした音声認識装置に用いられる辞書を作成する音声認識用辞書作成装置およびその辞書を用いて音声を認識する音声認識装置等として、特に、省略語を含む語彙を認識する音声認識装置等として、例えば、デジタル放送受信機やカーナビゲーション装置等として利用することができる。  
15

## 請 求 の 範 囲

1. 音声認識用辞書を作成する音声認識用辞書作成装置であって、
  - 1 以上の中から構成される認識対象語について、発声のし易さを考慮したルールに基づいて、前記認識対象語の省略語を生成する省略語生成手段と、

生成された省略語を前記認識対象語とともに前記音声認識用辞書として記憶する語彙記憶手段と  
を備えることを特徴とする音声認識用辞書作成装置。
- 10 2. 前記音声認識用辞書作成装置はさらに、
  - 前記認識対象語を構成単語に分割する単語分割手段と、

分割された構成単語ごとの読みに基づいて、構成単語ごとのモーラ列を生成するモーラ列生成手段とを備え、

前記省略語生成手段は、前記モーラ列生成手段によって生成された構成単語ごとのモーラ列に基づいて、構成単語ごとのモーラ列からモーラを取り出して連接することにより、1個以上のモーラからなる省略語を生成する

ことを特徴とする請求の範囲1記載の音声認識用辞書作成装置。
- 20 3. 前記省略語生成手段は、
  - モーラを用いた省略語の生成規則を格納している省略語生成規則格納部と、

前記構成単語ごとのモーラ列からモーラを取り出して連接することにより、1個以上のモーラからなる省略語の候補を生成する候補生成部と、

生成された省略語の候補に対して、前記省略語生成規則格納部に格納された生成規則を適用することで、最終的に生成する省略語を決定する

省略語決定部とを有する

ことを特徴とする請求の範囲 2 記載の音声認識用辞書作成装置。

4. 前記省略語生成規則格納部には、複数の生成規則が格納され、

5 前記省略語決定部は、生成された省略語の候補について、前記省略語生成規則格納部に格納された複数の規則それぞれに対する尤度を算出し、算出した尤度を総合的に勘案することによって発声確率を決定し、

前記語彙記憶手段は、前記省略語決定部によって決定された省略語および発声確率を前記認識対象語とともに記憶する

10 ことを特徴とする請求の範囲 3 記載の音声認識用辞書作成装置。

5. 前記省略語決定部は、前記複数の規則それぞれに対する尤度に、対応する重み付け係数を乗じて得られる値を合計することによって前記発声確率を決定する

15 ことを特徴とする請求の範囲 4 記載の音声認識用辞書作成装置。

6. 前記省略語決定部は、前記省略語の候補に対する発声確率が一定のしきい値を超える場合に、最終的に生成する省略語と決定する

ことを特徴とする請求の範囲 5 記載の音声認識用辞書作成装置。

20

7. 前記省略語生成規則格納部には、単語の係り受けに関する第 1 の規則が格納され、

前記省略語決定部は、前記第 1 の規則に基づいて、前記候補の中から最終的に生成する省略語を決定する

25 ことを特徴とする請求の範囲 4 記載の音声認識用辞書作成装置。

8. 前記第1の規則には、修飾語と被修飾語とを対にすることによって省略語を生成するという条件が含まれる  
ことを特徴とする請求の範囲7記載の音声認識用辞書作成装置。

5 9. 前記第1の規則には、省略語を構成する修飾語と被修飾語との距離と前記尤度との関係を示す規則が含まれる  
ことを特徴とする請求の範囲7記載の音声認識用辞書作成装置。

10. 前記省略語生成規則格納部には、省略語を生成するときに構成単語のモーラ列から取り出される部分モーラ列の長さおよび構成単語における位置の少なくとも1つに関する第2の規則が格納され、  
前記省略語決定部は、前記第2の規則に基づいて、前記候補の中から最終的に生成する省略語を決定する  
ことを特徴とする請求の範囲4記載の音声認識用辞書作成装置。

15

11. 前記第2の規則には、前記部分モーラ列の長さを示すモーラ数と前記尤度との関係を示す規則が含まれる  
ことを特徴とする請求の範囲10記載の音声認識用辞書作成装置。

20 12. 前記第2の規則には、前記部分モーラ列の構成単語における位置を示す構成単語の先頭からの距離に対応するモーラ数と前記尤度との関係を示す規則が含まれる  
ことを特徴とする請求の範囲10記載の音声認識用辞書作成装置。

25 13. 前記省略語生成規則格納部には、省略語を構成する部分モーラ列の連なりに関する第3の規則が格納され、

前記省略語決定部は、前記第3の規則に基づいて、前記候補の中から最終的に生成する省略語を決定する  
ことを特徴とする請求の範囲4記載の音声認識用辞書作成装置。

5 14. 前記第3の規則には、連接された2つの部分モーラ列における前に位置する部分モーラ列の最後のモーラと後に位置する部分モーラ列の先頭のモーラとの組み合わせと前記尤度との関係を示す規則が含まれることを特徴とする請求の範囲13記載の音声認識用辞書作成装置。

10 15. 前記音声認識用辞書作成装置は、さらに、  
認識対象語を含んだ文字列情報から認識対象語を抽出する条件を格納  
している抽出条件格納手段と、  
認識対象語を含んだ文字列情報を取得する文字列情報取得手段と、  
前記抽出条件格納手段に格納されている条件に従って、前記文字列情  
15 報取得手段によって取得された文字列情報から認識対象語を抽出し、前  
記単語分割手段に送出する認識対象語抽出手段とを備える  
ことを特徴とする請求の範囲2記載の音声認識用辞書作成装置。

16. 入力された音声を、音声認識用辞書に登録されている語彙に対応  
20 するモデルによって照合を行って認識する音声認識装置であって、  
請求の範囲1記載の音声認識用辞書作成装置によって作成された音声  
認識用辞書を用いて前記音声を認識する認識手段を備える  
ことを特徴とする音声認識装置。

25 17. 前記音声認識用辞書には、前記省略語と当該省略語の発声確率と  
が前記認識対象語とともに登録され、

前記認識手段は、前記音声認識用辞書に登録されている発声確率を考慮して前記音声の認識を行うことを特徴とする請求の範囲 1 6 記載の音声認識装置。

5 18. 前記認識手段は、前記音声の認識結果である候補とともに当該候補の尤度を生成し、生成した尤度に前記発声確率に対応する尤度を加算し、得られた加算値に基づいて前記候補を最終的な認識結果として出力する

ことを特徴とする請求の範囲 1 7 記載の音声認識装置。

10

19. 前記音声認識装置は、さらに、

前記音声に対して認識した省略語と当該省略語に対応する認識対象語とを使用履歴情報として格納する省略語使用履歴格納手段と、

15 前記省略語使用履歴格納手段に格納された使用履歴情報に基づいて、前記省略語生成手段による省略語の生成を制御する省略語生成制御手段とを備える

ことを特徴とする請求の範囲 1 6 記載の音声認識装置。

20. 前記音声認識用辞書作成装置の省略語生成手段は、

モーラを用いた省略語の生成規則を格納している省略語生成規則格納部と、

前記構成単語ごとのモーラ列からモーラを取り出して連接することにより、1個以上のモーラからなる省略語の候補を生成する候補生成部と、

25 生成された省略語の候補に対して、前記省略語生成規則格納部に格納された生成規則を適用することで、最終的に生成する省略語を決定する省略語決定部とを有し、

前記省略語生成制御手段は、前記省略語生成規則格納部に格納される生成規則を変更、削除または追加することによって前記省略語の生成を制御する

ことを特徴とする請求の範囲 1 9 記載の音声認識装置。

5

2 1. 前記音声認識装置は、さらに、

前記音声に対して認識した省略語と当該省略語に対応する認識対象語とを使用履歴情報として格納する省略語使用履歴格納手段と、

前記省略語使用履歴格納手段に格納された使用履歴情報に基づいて、

10 前記音声認識用辞書に格納されている省略語に対する編集を行う辞書編集手段とを備える

ことを特徴とする請求の範囲 1 6 記載の音声認識装置。

2 2. 前記音声認識用辞書には、前記省略語と当該省略語の発声確率と  
15 が前記認識対象語とともに登録され、

前記辞書更新手段は、前記省略語の発声確率を変更することによって前記省略語に対する編集を行う

ことを特徴とする請求の範囲 2 1 記載の音声認識装置。

20 2 3. 入力された音声を、音声認識用辞書に登録されている語彙に対応するモデルによって照合を行って認識する音声認識装置であって、

請求の範囲 1 記載の音声認識用辞書作成装置と、

前記音声認識用辞書作成装置によって作成された音声認識用辞書を用いて前記音声を認識する認識手段と

25 を備えることを特徴とする音声認識装置。

24. 音声認識用辞書を作成する音声認識用辞書作成方法であって、

1 以上の単語から構成される認識対象語について、発声のし易さを考慮したルールに基づいて、前記認識対象語の省略語を生成する省略語生成ステップと、

- 5 生成された省略語を前記認識対象語とともに前記音声認識用辞書に登録する語彙登録ステップと  
含むことを特徴とする音声認識用辞書作成方法。

25. 前記音声認識用辞書作成方法はさらに、

- 10 前記認識対象語を構成単語に分割する単語分割ステップと、  
分割された構成単語ごとの読みに基づいて、構成単語ごとのモーラ列を生成するモーラ列生成ステップとを含み、  
前記省略語生成ステップでは、前記モーラ列生成手段によって生成された構成単語ごとのモーラ列に基づいて、構成単語ごとのモーラ列から  
15 モーラを取り出して連接することにより、1個以上のモーラからなる省略語を生成する  
ことを特徴とする請求の範囲24記載の音声認識用辞書作成方法。

26. 入力された音声を、音声認識用辞書に登録されている語彙に対応

- 20 するモデルによって照合を行って認識する音声認識方法であって、  
請求の範囲24記載の音声認識用辞書作成方法によって作成された音声認識用辞書を用いて前記音声を認識する認識ステップを含む  
ことを特徴とする音声認識方法。

- 25 27. 入力された音声を、音声認識用辞書に登録されている語彙に対応するモデルによって照合を行って認識する音声認識方法であって、

請求の範囲 2 4 記載の音声認識用辞書作成方法におけるステップと、  
前記音声認識用辞書作成方法によって作成された音声認識用辞書を用  
いて前記音声を認識するステップとを  
含むことを特徴とする音声認識方法。

5

2 8. 音声認識用辞書を作成する音声認識用辞書作成装置のためのプロ  
グラムであって、

請求の範囲 2 4 記載の音声認識用辞書作成方法におけるステップをコ  
ンピュータに実行させる

10 ことを特徴とするプログラム。

2 9. 入力された音声を、音声認識用辞書に登録されている語彙に対応  
するモデルによって照合を行って認識する音声認識装置のためのプログ  
ラムであって、

15 請求の範囲 2 6 記載の音声認識方法におけるステップをコンピュータ  
に実行させる  
ことを特徴とするプログラム。

図1

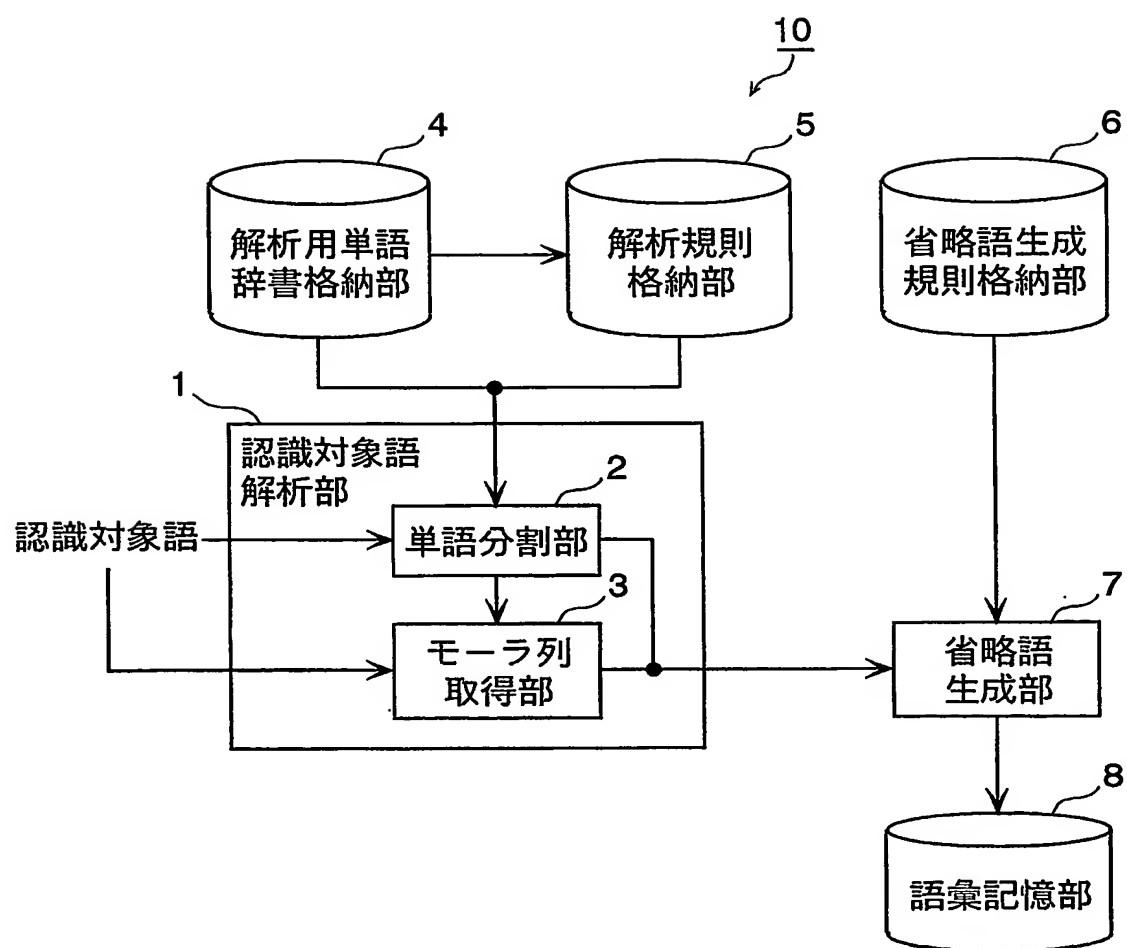


図2

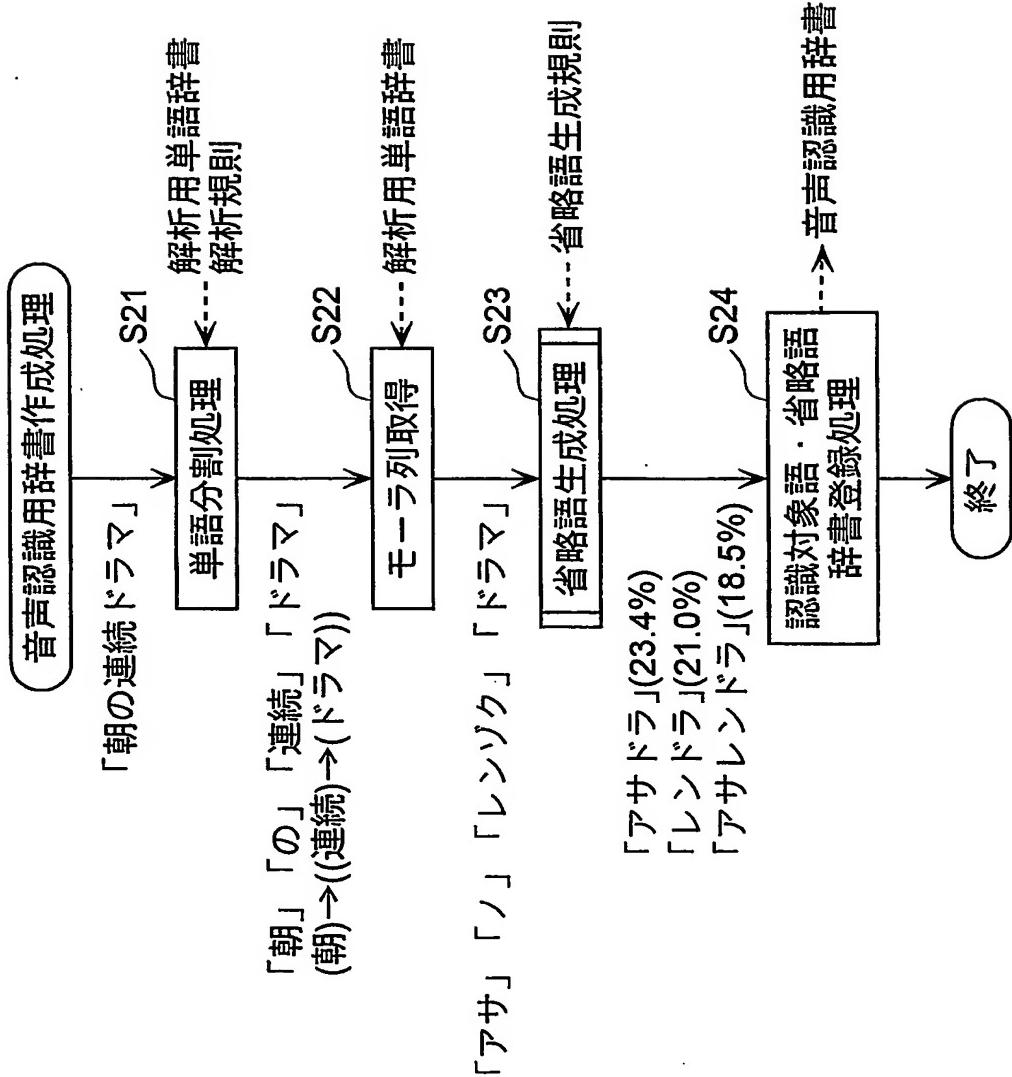


図3

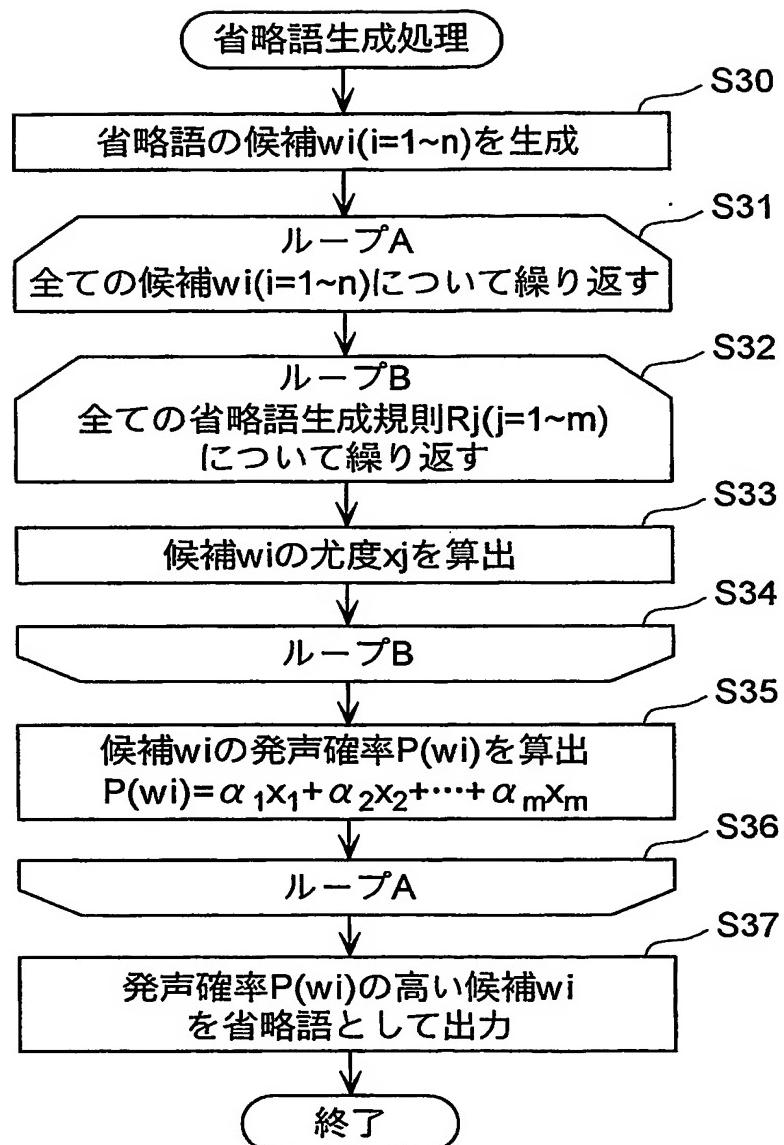


図4

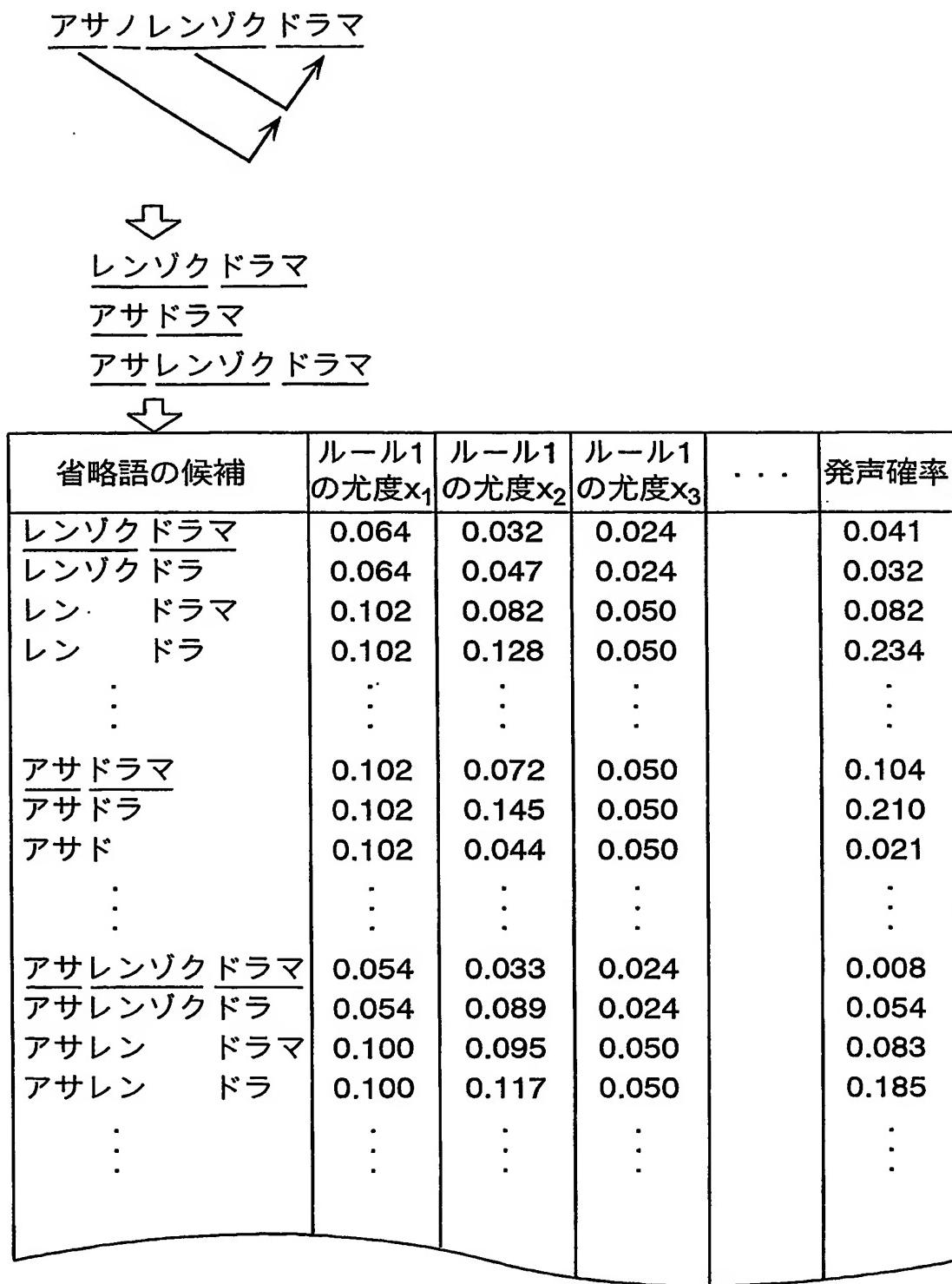


図5

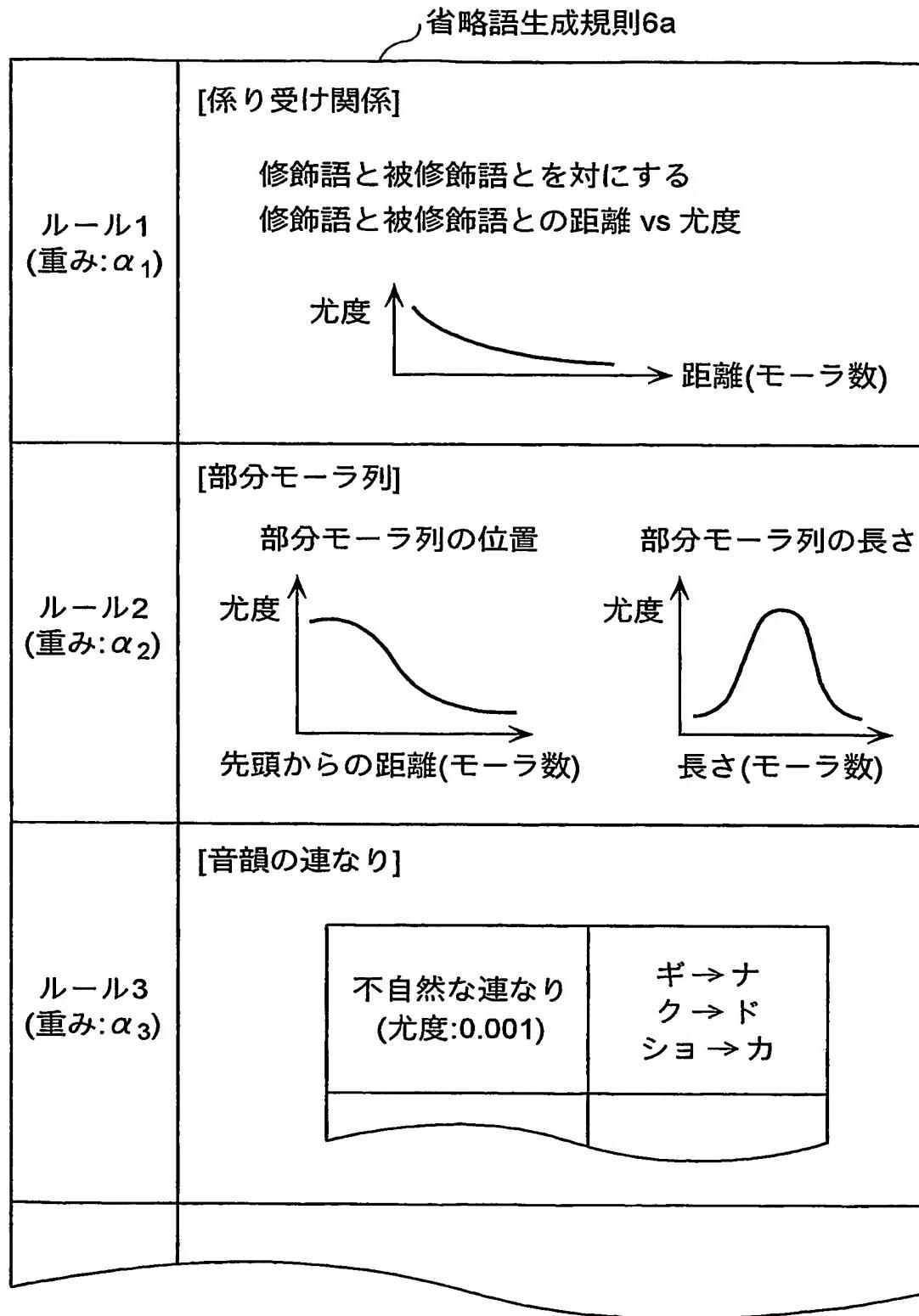


図6

音声認識用辞書8a

認識対象語	省略語	発声確率
「朝の連続 ドラマ」	レンドラ	0.234
	アサドラ	0.210
	アサレンドラ	0.185
⋮	⋮	⋮

四七

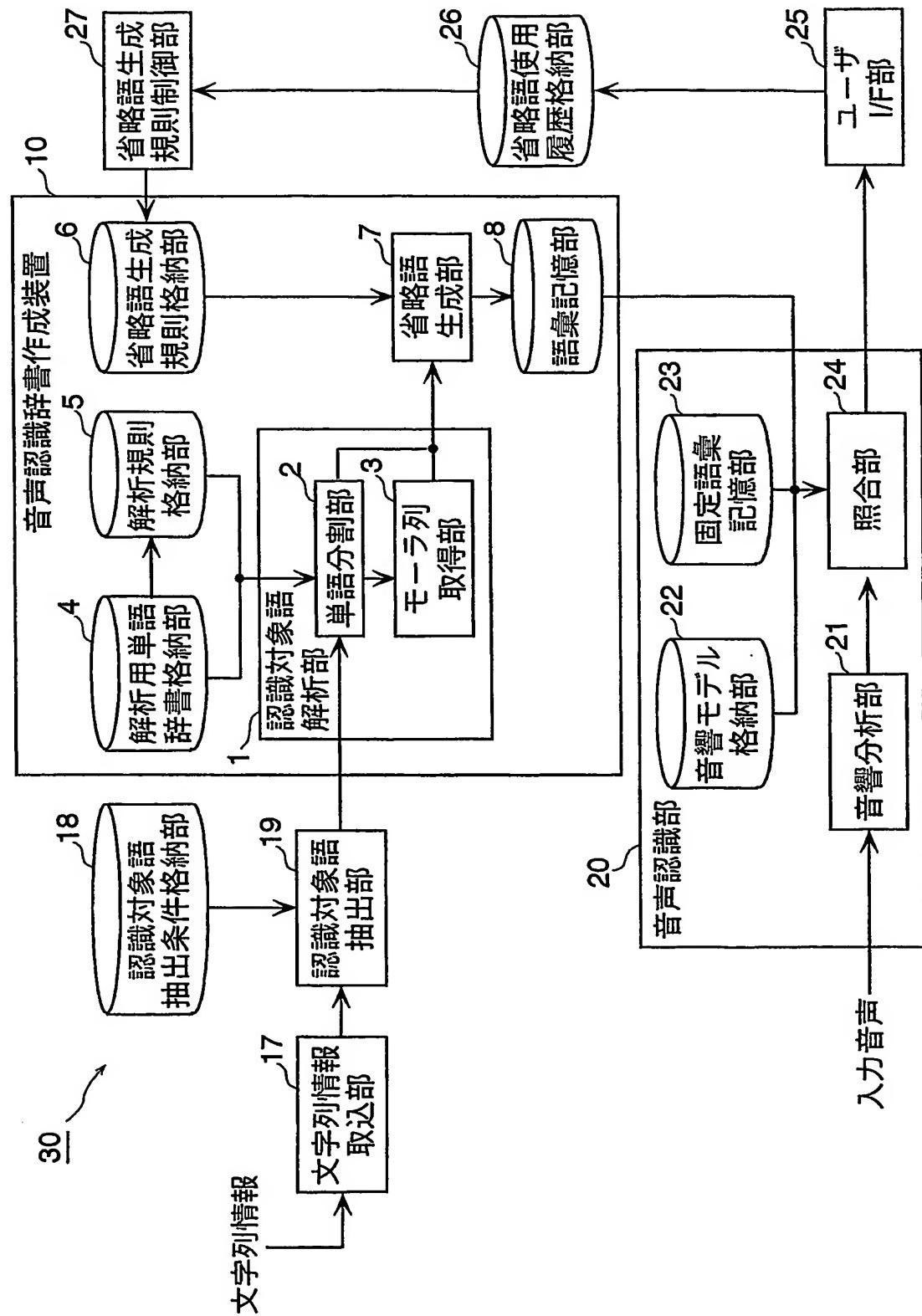


図8

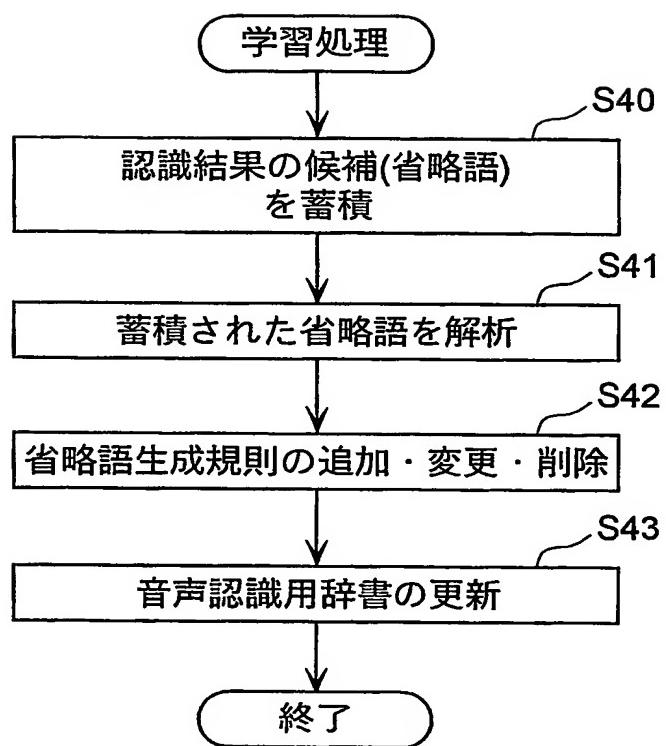


図9

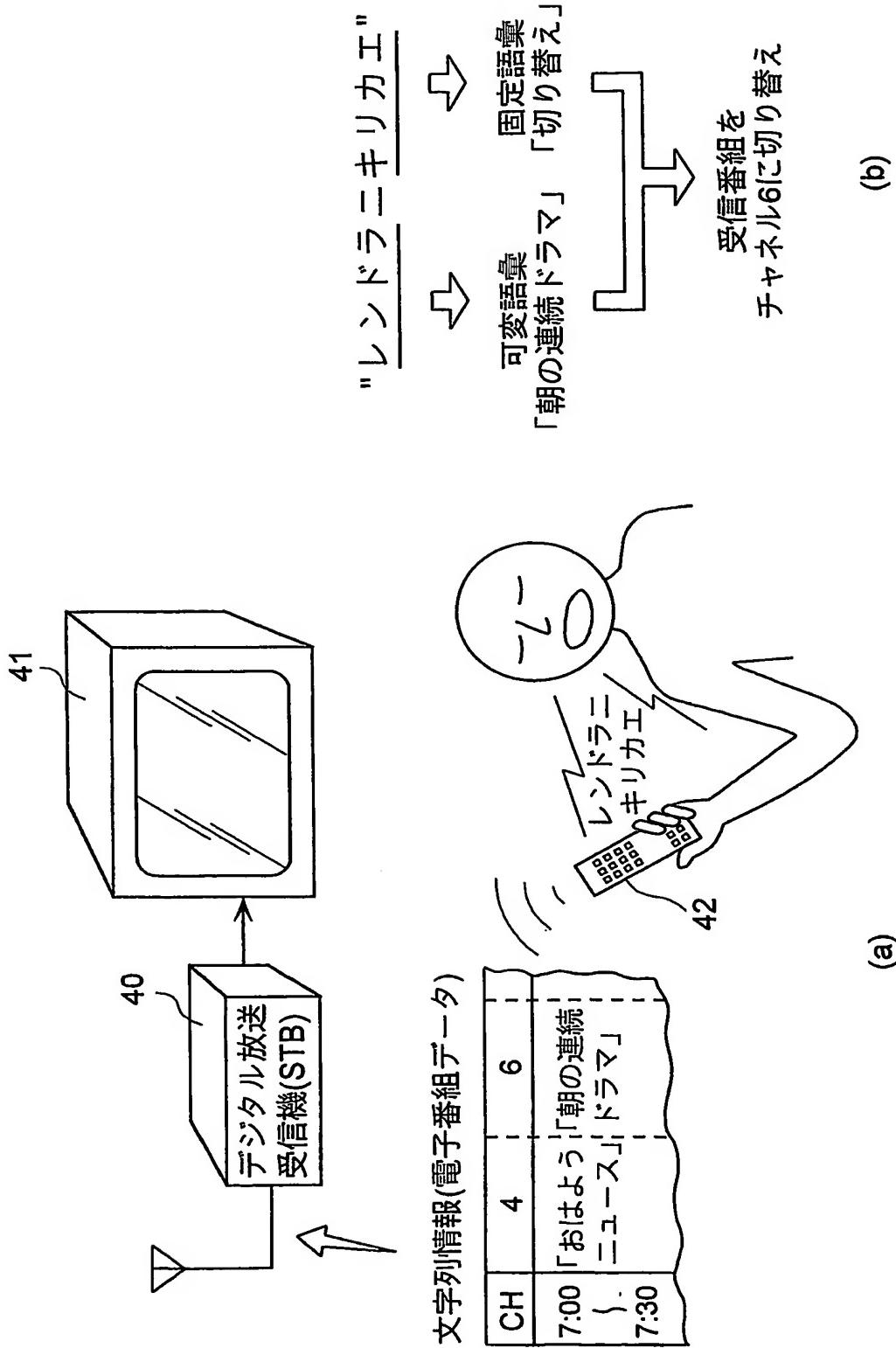


図10

認識対象語	省略語
全国人民代表大会	全人代
北京大学	北大
清華大学	清大
公共汽車	公車
:	:

(a)

認識対象語	省略語
Electric intelligence	ELINT
Elizabeth	Eliz Eliza
Eleanor	Ella
Frankfurter	Frank
Coop for American Remittance Europe	CARE
James	Jim
Joseph	Jock
Joseph	Joe
Kitten	Kit
Mississippi	Miss
Tactical an navigation	TACAN
:	:

(b)

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/14168

**A. CLASSIFICATION OF SUBJECT MATTER**

Int.Cl<sup>7</sup> G10L15/06, G10L15/18

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G10L15/06, G10L15/18, G06F3/16, G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Toroku Jitsuyo Shinan Koho	1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
JICST FILE (JOIS)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 3-194653 A (Tokai Television Broadcasting Co., Ltd.), 26 August, 1991 (26.08.91), Full text; all drawings (Family: none)	1-6, 10-12, 15-18, 23-29
Y	JP 2002-41081 A (Sharp Corp.), 08 February, 2002 (08.02.02), Full text; all drawings (Family: none)	1-6, 10-12, 15-18, 23-29
A	JP 2001-34290 A (Omron Corp.), 09 February, 2001 (09.02.01), Full text; all drawings (Family: none)	1-29

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search  
26 January, 2004 (26.01.04)

Date of mailing of the international search report  
10 February, 2004 (10.02.04)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP03/14168

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 8-272789 A (Mitsubishi Electric Corp.), 18 October, 1996 (18.10.96), Full text; all drawings (Family: none)	1-15
A	JP 11-328166 A (Brother Industries, Ltd.), 30 November, 1999 (30.11.99), Full text; all drawings (Family: none)	1-15
A	JP 11-110408 A (Sharp Corp.), 23 April, 1999 (23.04.99), Full text; all drawings (Family: none)	1-15

## A. 発明の属する分野の分類（国際特許分類（IPC））

Int. C17 G10L15/06, G10L15/18

## B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int. C17 G10L15/06, G10L15/18,  
G06F3/16, G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996
日本国公開実用新案公報	1971-2004
日本国実用新案登録公報	1996-2004
日本国登録実用新案公報	1994-2004

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

JICSTファイル（JOIS）

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 3-194653 A (東海テレビ放送株式会社) 1991.08.26, 全文, 全図 (ファミリーなし)	1-6, 10-12, 15-18, 23-29
Y	JP 2002-41081 A (シャープ株式会社) 2002.02.08, 全文, 全図 (ファミリーなし)	1-6, 10-12, 15-18, 23-29

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

## の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

## 国際調査を完了した日

26.01.2004

## 国際調査報告の発送日

10.2.2004

## 国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

## 特許庁審査官（権限のある職員）

樋本 剛

5C 3352

電話番号 03-3581-1101 内線 3541

C(続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2001-34290 A (オムロン株式会社) 2001. 02. 09, 全文, 全図 (ファミリーなし)	1-29
A	JP 8-272789 A (三菱電機株式会社) 1996. 10. 18, 全文, 全図 (ファミリーなし)	1-15
A	JP 11-328166 A (プラザー工業株式会社) 1999. 11. 30, 全文, 全図 (ファミリーなし)	1-15
A	JP 11-110408 A (シャープ株式会社) 1999. 04. 23, 全文, 全図 (ファミリーなし)	1-15